

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

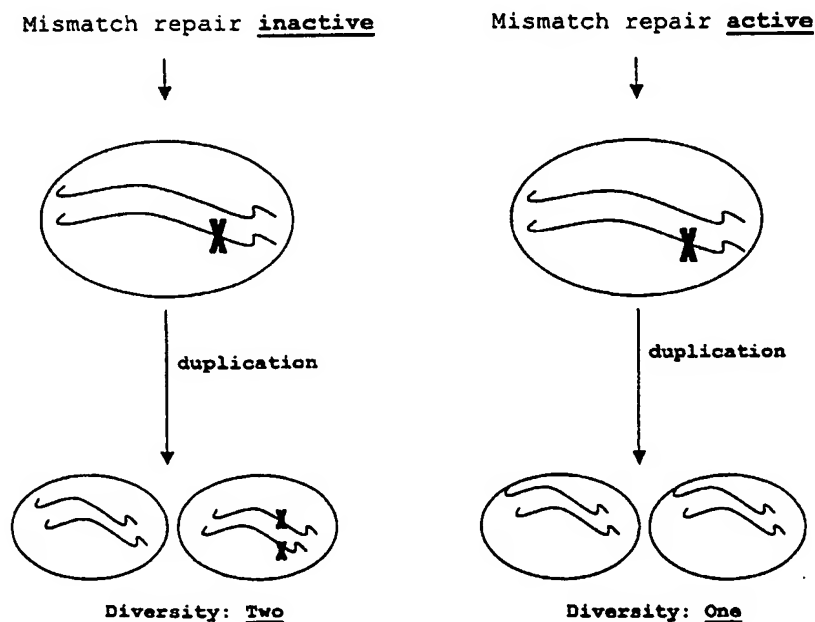
**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>7</sup> : C12N 1/15, 15/10, C12P 9/00 // (C12N 1/15, C12R 1:66) (C12N 1/15, C12R 1:77)</p>	A1	<p>(11) International Publication Number: <b>WO 00/50567</b></p> <p>(43) International Publication Date: 31 August 2000 (31.08.00)</p>
<p>(21) International Application Number: PCT/DK00/00063</p> <p>(22) International Filing Date: 17 February 2000 (17.02.00)</p> <p>(30) Priority Data: PA 1999 00253 24 February 1999 (24.02.99) DK</p> <p>(71) Applicant: NOVO NORDISK A/S [DK/DK]; Enzyme Business Patents, Novo Alle, DK-2880 Bagsvaerd (DK).</p> <p>(72) Inventors: BORCHERT, Torben, Vedel; Vordingborggade 6a, 31, DK-2100 Copenhagen Ø (DK). CHRISTIANSEN, Lars (deceased). VIND, Jesper; Bagsværdvej 115, DK-2800 Lyngby (DK).</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report.</p>

(54) Title: FUNGAL CELLS WITH INACTIVATED DNA MISMATCH REPAIR SYSTEM

## (57) Abstract

A process for making DNA libraries in filamentous fungal cells using a novel cloned gene involved in the mismatch repair system of filamentous fungal cells.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## FUNGAL CELLS WITH INACTIVATED DNA MISMATCH REPAIR SYSTEM

## 5 FIELD OF INVENTION

A process for making DNA libraries in filamentous fungal cells using a novel cloned gene involved in the mismatch repair system of filamentous fungal cells.

## 10 BACKGROUND OF THE INVENTION

The mismatch repair system is a system within cells which recognises mismatches in newly synthesised duplex DNA sequences.

The mismatch repair system then either corrects the  
15 mismatches which are seen as errors by e.g. using the methylated "old" strain as template or alternatively it may mediate degradation of the duplex DNA sequences which comprise the mismatches.

Independently on the precise mechanism the end result  
20 will be that the "mismatch repair system" will limit the "diversity" within a cell, diversity being represented as duplex DNA sequences which comprise mismatches.

For example a duplex DNA sequence which comprises a single mismatch represents a diversity of two different DNA  
25 sequences within the cell. If the mismatch repair system corrects the mismatch there will only be a diversity of one within the cell.

Alternatively, if the mismatch repair system mediates the degradation of such a duplex DNA sequence the diversity will be  
30 lost. See figure 1 for a graphic illustration on how the mismatch repair system may work within a cell.

Consequently, if duplex DNA sequences comprising mismatches represent a DNA library of interest, then the diversity of this library may be limited when transformed  
35 (placed) into cells with an active mismatch repair system.

The art provides a solution to this problem by making cells wherein the mismatch system is inactive.

EP 449923 describes bacterial cells wherein the mismatch system is inactivated.

WO 97/37011 describes yeast cells wherein the mismatch system is inactivated. See the working examples of this document.

WO 97/05268 describes mice cells wherein the mismatch system is inactivated. See the working examples of this document.

#### 10 SUMMARY OF THE INVENTION:

The problem to be solved by the present invention is to provide an improved strategy for making DNA libraries in filamentous fungal cells. A filamentous fungal cell population comprising such a DNA library may then be used to select a polypeptide of interest. Also polynucleotide sequences with particular properties might be selected, such as promoters, terminators and other regulatory elements with changed/improved properties.

The solution is based on that the present inventors have cloned a NOVEL gene involved in the mismatch repair system of a filamentous fungal cell. Further, this gene is the first gene cloned which is involved in the mismatch repair system of a filamentous fungal cell.

By inactivating this gene in a filamentous cell it is possible to obtain a filamentous cell which is deficient in its mismatch repair system and which is highly useful for preparing DNA libraries in filamentous fungal cells.

The gene comprises a very characterising DNA sequence encoding the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2.

This DNA has been used to clone the full length gene encoding the polypeptide sequence shown in positions 1-940 of SEQ ID NO 2. See working examples herein (vide infra).

The gene cloned as described in working examples herein is a gene cloned from an *Aspergillus oryzae* filamentous fungal cell.

However, based on the novel sequence information provided herein it is routine work for the skilled person to clone

similar homologous genes from other filamentous fungal cells by e.g. standard hybridisation or PCR technology, preferably by using the DNA sequence encoding the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2 as a basis for making a hybridisation probe or PCR primers.

Accordingly, in a first aspect the present invention relates to a filamentous fungal cell, wherein a gene involved in the mismatch repair system has been inactivated and in which the gene involved in the mismatch repair system comprises:

- 10 (a) a DNA sequence encoding the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2; or
- (b) a DNA sequence encoding a polypeptide sequence which is at least 70% identical to the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2;

15 and

in a second aspect the present invention relates to a filamentous fungal cell, wherein a gene involved in the mismatch repair system has been inactivated and in which the gene involved in the mismatch repair system comprises:

- 20 (a) a DNA sequence encoding the polypeptide sequence shown in positions 1-940 of SEQ ID NO 2; or
- (b) a DNA sequence encoding a polypeptide sequence which is at least 70% identical to the polypeptide sequence shown in positions 1-940 of SEQ ID NO 2.

25

As stated above a filamentous fungal cell of the first or second aspect of the invention is very suitable for making a DNA library of interest in filamentous fungal cells.

Accordingly, in a third aspect the present invention relates to a process for preparing a filamentous fungal cell population wherein individual cells in the population comprise individually different DNA sequences of interest representing a DNA library of interest comprising following steps:

- 35 (a) placing individually different DNA sequences of interest in a filamentous fungal cell population comprising a filamentous fungal cell of the first or second aspect of the invention; and
- (b) growing the population of (a) for a period of time allowing an individual DNA sequence of interest in the

population to be duplicated at least once under conditions wherein the mismatch repair system gene of the first or second aspect of the invention has been inactivated.

5

One of the advantages of allowing an individual mismatch repair inactivated filamentous fungal cell duplicated DNA of interest at least once as described under step (b) of the third aspect is illustrated in figure 1. As can be seen in figure 1  
10 the process of the third aspect using a filamentous fungal mismatch repair inactivated cell as described herein allows preparation of a DNA library wherein eventual hetero duplex DNA mismatches are not corrected. This gives a DNA library with a higher diversity as compared to a DNA library made in a filamentous fungal cell NOT having an inactivated mismatch repair  
15 system (see figure 1). Duplication of DNA sequence of interest means that the two strands are replicated such that two separate sets of double stranded DNA are generated, each being based on a separate one of the two original strands.

20 A filamentous fungal cell population wherein individual cells in the population comprise a DNA library of interest as described above may be used to select a polypeptide of interest.

Accordingly, in a fourth aspect the present invention  
25 relates to a process for production of a polypeptide of interest comprising the steps of the third aspect and wherein the DNA sequences of interest encode a polypeptide of interest and which further comprises following step:

(c) selecting from the resultant population of  
30 filamentous fungal cells of step (b) of the third aspect a desired polypeptide of interest.

An advantage of the process of the fourth aspect may be that the polypeptide of interest is selected from a filamentous fungal cell expressing the polypeptide. Consequently, it is  
35 directly known that the polypeptide can be expressed from a filamentous fungal cell, which is usefully if it is subsequently required to produce the polypeptide in large scale in a filamentous fungal cell. This may be of particular

interest when the DNA library encodes polypeptides of interest which are derived from filamentous fungal cells, since it is known that filamentous fungal polypeptides preferably are produced in industrial relevant high yields in filamentous  
5 fungal cells.

This is contrary to a similar selection process using e.g. a yeast cell. Here the only thing known is that the selected polypeptide is capable of being expressed in yeast and later expression a filamentous fungal cell might give problems,  
10 especially if high yields are required.

#### DEFINITIONS:

Following section provides definitions of technical features in above-mentioned aspects of the invention.

15 The term "a gene" denotes herein a gene (a DNA sequence) will is capable of being expressed into a polypeptide within said cell. Accordingly, said gene sequence will be defined as an open reading frame starting from a start codon (normally "ATG", "GTG", or "TTG") and ending at a stop codon  
20 (normally "TAA", TAG" or "TGA").

In order to express said gene there must be elements, as known in the art, in connection with the gene, necessary for expression of the gene within the cell. Such standard elements may include a promoter, a ribosomal binding site, a termination  
25 sequence, and may be others elements as known in the art.

The term "mismatch repair system" shall herein be understood according to the art, as a system within cells which recognises mismatches in duplex DNA sequences. See e.g. WO 97/37011, page 1, line 21-28)

30 The mismatch repair system then either corrects the mismatches which are seen as errors by e.g. using the methylated "old" strain as template or alternatively it may mediate degradation of the duplex DNA sequences which comprise the mismatches.

35 Independently on the precise mechanism the end result will be that the "mismatch repair system" will limit the "diversity" within the cell represented by such duplex DNA sequences which comprise mismatches.



For example a duplex DNA sequence which comprises a single mismatch represents a diversity of two different DNA sequences within the cell. If the mismatch repair system corrects the mismatch there will only be a diversity of one within the cell. Alternatively, if the mismatch repair system mediates the degradation of such a duplex DNA sequence this diversity will be lost.

A polypeptide encoded by a gene involved in the mismatch repair system recognises a mismatch by a mechanism involving binding to the mismatch.

Accordingly, a suitable assay to test whether or not a filamentous fungal cell as described herein is inactivated in its mismatch repair system is to use a "gel shift assay" or alternatively termed a "gel retardation assay". This is a standard assay used in the art. See WO 97/05268, page 16,17 and 25.

The principle in such an assay is that cell extracts are prepared of both (a) a filamentous fungal cell wherein the gene, as described herein, involved in the mismatch repair system is inactivated; and (b) the corresponding filamentous fungal cell wherein the gene is NOT inactivated. These extracts are then bound/mixed to oligonucleotides containing the base-pair mismatched G:T; G:A; G:G; A:C, and an extrahelical TG dinucleotide and run on a nondenaturing gel.

If the gel shift assay demonstrates that the control filamentous fungal cell wherein the gene is NOT inactivated comprises any protein(s) which binds to any of above mentioned oligonucleotides and these binding protein(s) are NOT seen in the filamentous fungal cell wherein the gene, as described herein, involved in the mismatch repair system is inactivated then it is a confirmation that the mismatch repair system in the latter is inactivated.

A detailed description of a suitable gel shift assay is provided in working example 1 herein.

The sequence identity in relation to the terms

"a DNA sequence encoding a polypeptide sequence which is at least 70% identical to the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2" and

``a DNA sequence encoding a polypeptide sequence which is at least 70% identical to the polypeptide sequence shown in positions 1-940 of SEQ ID NO 2'';

is determined as the degree of identity between two  
5 sequences indicating a derivation of the first sequence from the second. The identity may suitably be determined by means of computer programs known in the art such as GAP provided in the GCG program package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science  
10 Drive, Madison, Wisconsin, USA 53711) (Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453). Using GAP with the following settings for polypeptide sequence comparison: GAP creation penalty of 3.0 and GAP extension penalty of 0.1, the polypeptide encoded by an  
15 analogous DNA sequence of the invention exhibits a degree of identity preferably of at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and especially at least 97% with amino acid sequence shown in positions 683-758 of SEQ ID NO 2, according to the first  
20 aspect of the invention; or with amino acid sequence shown in positions 1-940 of SEQ ID NO 2, according to the second aspect of the invention.

The term ``DNA library'' denotes herein a library of at least two different DNA sequences. For many practical purposes  
25 the library is must bigger. Accordingly, the DNA library preferably comprises at least 1000 different DNA sequences, more preferably at least 10000 different DNA sequences, and even more preferably at least 100000 different DNA sequences.

The term ``placing individually different DNA sequences of  
30 interest in a filamentous fungal cell population'' in relation to step (a) in the process of the third aspect of the invention shall herein be understood broadly in the sense that it is NOT identical DNA sequences of interest which are placed in the filamentous fungal cell population. In the present context,  
35 relating to a process for making a DNA library using a mismatch repair deficient cell, the term should preferably denotes a situation wherein a cell within the filamentous fungal cell population comprises at least two different DNA sequences of

interest which are so partially homologous that they are capable of hybridising/recombining to each other within the cell. It is within the skilled persons general knowledge to determine how partially homologous such sequences have to be in order to obtain said recombination within the cell.

A practical example may be that single stranded oligonucleotide sequences partially homologous to chromosomal DNA sequence are placed within the cell or duplex DNA sequences comprising mismatches (e.g. comprised within a vector) are placed within the cell. See below for further description of such examples.

The specific experimental way of placing these DNA sequences within a filamentous cell may be done according to any of the many suitable techniques, such as transformation techniques.

The term "growing the population of (a) for a period of time allowing an individual DNA sequence of interest in the population to be duplicated at least once under conditions wherein the mismatch repair system gene has been inactivated" according to step (b) of the third aspect of the invention denotes that after an individual cell has duplicated itself at least once then the mismatch repair system may be activated again without losing the advantage of the process. The technical reason for this is illustrated in figure 1. In this example a duplex DNA sequence comprising a single mismatch is placed in filamentous cell. After the cell has been duplicated once under conditions wherein the mismatch repair system gene has been inactivated the two individually different single stranded DNA sequences within the duplex DNA have individually been duplicated providing two different duplex sequences, one in each duplicated cell, without any mismatches. Accordingly, since such a cell does NOT comprise duplex DNA sequences of interest having mismatches then there is no technical need to maintaining the mismatch repair system inactivated.

In sections below are described preferred embodiments of the invention by way of examples only.

DRAWINGS:

Figure 1:

This figure illustrates an example wherein a duplex DNA sequence comprising a single mismatch is placed in filamentous cell. After the cell has been duplicated once under conditions wherein the mismatch repair system gene has been inactivated the two individually different single stranded DNA sequences within the duplex DNA have individually been duplicated providing two different duplex sequences, one in each duplicated cell, without any mismatches. On the contrary, in a cell wherein the mismatch repair system is active, a mismatch within a duplex is corrected.

Figure 2:

This figure shows three partial *Aspergillus oryzae* polypeptide sequences: 'msh2'Ao-coll0/13/15; derived from cloned PCR fragments. The three partial polypeptide sequences are aligned with two other partial polypeptide sequences of known mismatch repair proteins: a human mismatch repair protein, msh2-human.p; and a fungal *Saccharomyces cerevisiae* mismatch repair protein, S.c. msh2. The underlined sequences in the figure derive from the construction of the PCR fragments.

Figure 3:

This figure shows an alignment of the proposed polypeptide sequence of the putative *Aspergillus oryzae* mismatch repair protein (Ao.MSH2) with the polypeptide sequences of three known mismatch repair proteins from human (msh2-human.p), mouse (msh2-mus.p), and yeast (S.c. msh2).

30

## DETAILED DESCRIPTION OF THE INVENTION

A filamentous fungal cell, as described herein, wherein a gene, as described herein, involved in the mismatch repair system has been inactivated.

35

Inactivation of a gene involved in the mismatch repair system:

The NOVEL gene, as described herein, involved in the mismatch repair system may be inactivated by any of the numerous known techniques known to the skilled person.

An embodiment of the invention relates to a filamentous  
5 fungal cell as described herein, wherein the gene involved in the mismatch repair is defective.

Numerous methods are known to the skilled person to make a gene defective when the DNA sequence is KNOWN. These methods includes deleting part of the DNA sequence of the gene;  
10 introducing frame-shift mutations by deleting or inserting nucleotides; introducing stop codons etc.

A preferred embodiment of the invention relates to a filamentous fungal cell as described herein, wherein the gene involved in the mismatch repair has been inactivated  
15 transitorily.

Similarly to above, a number of methods are known to the skilled person for doing this, including insertion of a regulable promoter upstream of the gene or e.g. permanently deleting part of the gene on the chromosome followed by  
20 inserting a vector (e.g. a plasmid) into the cell which comprises the gene. The plasmid may then comprise a regulable promoter up-stream of the gene or it may be that the plasmid can be removed from the cell when the mismatch repair system shall be inactivated transitorily and then re-inserted into the cell  
25 when the mismatch repair system shall be re-activated.

It is within the skilled persons general knowledge to choose the appropriate strategy for a specific technical purpose.

A preferred way to make a filamentous fungal cell which  
30 is capable of transitorily inactivate the mismatch repair system as described herein is first to permanently inactive the mismatch repair gene described herein on the chromosome of the cell followed by inserting a plasmid into the cell which comprises the gene, wherein the plasmid is characterised by  
35 that it comprises a suitable replication initiating sequence and a suitable selectable marker.

Preferably the suitable replication initiating sequence is AMA1 (Gems, D., et al. (1991, Gene 98:61-67).

A more detailed description of suitable replication initiating sequences and suitable selectable markers is provided immediately below and in working example 4 herein is provided an example of this strategy using a plasmid comprising 5 AMA1 as replication initiating sequence and AmdS as selectable marker.

#### Replication initiating sequences

As used herein, the term "fungal replication initiating 10 sequence" is defined as a nucleic acid sequence which is capable of supporting autonomous replication of an extrachromosomal molecule, e.g., a plasmid or a DNA vector, in a fungal host cell, normally without structural rearrangement of the plasmid or integration into the host cell genome. The replication ini- 15 tiating sequence may be of any origin as long as it is capable of mediating replication initiating activity in a fungal cell. Preferably, the replication initiating sequence is obtained from a filamentous fungal cell, more preferably a strain of *Aspergillus*, *Fusarium* or *Alternaria*, and even more preferably, a 20 strain of *A. nidulans*, *A. oryzae*, *A. niger*, *F. oxysporum* or *Alternaria alternata*.

A replication initiating sequence may be identified by methods well-known in the art. For instance, the sequence may be identified among genomic fragments derived from the organism 25 in question as a sequence capable of sustaining autonomous replication in yeast, (Ballance and Turner, *Gene*, 36 (1985), 321-331), an indication of a capability of autonomous replication in filamentous fungal cells. The replication initiating activity in fungi of a given sequence may also be determined by 30 transforming fungi with contemplated plasmid replicators and selecting for colonies having an irregular morphology, indicating loss of a sectorial plasmid which in turn would lead to lack of growth on selective medium when selecting for a gene found on the plasmid (Gems et al, *Gene*, 98 (1991) 61-67). AMA1 35 was isolated in this way. An alternative way to isolate a replication initiating sequence is to isolate natural occurring

plasmids (e.g. as disclosed by Tsuge et al., Genetics 146 (1997) 111-120 for *Alternaria aternata*).

Examples of replication initiating sequences include, but are not limited to, the ANS1 and AMA1 sequences of *Aspergillus* 5 *nidulans*, e.g., as described, respectively, by Cullen, D., et al. (1987, Nucleic Acids Res. 15:9163-9175) and Gems, D., et al. (1991, Gene 98:61-67).

The term "replication initiating activity" is used herein in its conventional meaning, i.e. to indicate that the sequence 10 is capable of supporting autonomous replication of an extrachromosomal molecule, such as a plasmid or a DNA vector in a fungal cell.

The term "without structural rearrangement of the plasmid" is used herein to mean that no part of the plasmid is deleted 15 or inserted into another part of the plasmid, nor is any host genomic DNA inserted into the plasmid.

#### Filamentous fungal selective marker

The term "selective pressure" is defined herein as culturing 20 a filamentous fungal cell, containing a DNA vector containing a fungal selective marker gene operably linked to a polynucleotide sequence of interest, in the presence of an effective amount or the absence of an appropriate selective agent. The effective amount of the selective agent is defined herein as an 25 amount sufficient for allowing the selection of cells containing the selection marker from cells which do not contain the selection marker.

In a preferred embodiment, the fungal selective marker is selected from the group of genes which encodes a product capable of providing resistance to biocide or viral toxicity, resistance to heavy metal toxicity, or prototrophy to auxotrophs. 30

In a more preferred embodiment, the prototrophy is obtained from an enzyme selected from the group of metabolic pathways consisting of nucleotide synthesis, cofactor synthesis, amino 35 acid synthesis, acetamide metabolism, proline metabolism, sulfate metabolism, and nitrate metabolism.

In an even more preferred embodiment, in the methods of the present invention the fungal selective marker is a gene selected from the group consisting of *argB* (ornithine carbamoyl-transferase), *amdS* (acetamidase), *bar* (phosphinothricin acetyl-  
5 transferase), *hemA* (5-aminolevulinate synthase), *hemB* (porphobilinogen synthase), *hygB* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *prn* (proline permease), *pyrG* (orotidine-5'-phosphate decarboxylase), *pyroA*, *riboB*, *sC* (sulfate adenylation transferase), and *trpC* (anthranilate synthase).

10 The fungal cell is cultivated in a suitable medium and under suitable conditions for screening or selecting for transformants harbouring the variant polynucleotide sequence of interest having or encoding the desired characteristic. The cultivation may be performed in accordance with methods well-known in  
15 the art for screening of polynucleotide variant libraries.

#### The filamentous fungal cell

The filamentous fungal cell as described herein includes all filamentous forms of the subdivision Eumycota and Oomycota.  
20 The filamentous fungi are characterised by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces*  
25 *cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative. In a preferred embodiment, the filamentous fungal cell is a cell of a species of, but is not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Scytalidium*,  
30 *Thielavia*, *Tolyposcladium*, and *Trichoderma*.

Examples of filamentous fungal cells of use in the present invention include an *Aspergillus* cell, an *Acremonium* cell, a *Fusarium* cell, a *Humicola* cell, a *Mucor* cell, a *Myceliophthora* cell, a *Neurospora* cell, a *Penicillium* cell, a *Thielavia* cell,  
35 a *Tolyposcladium* cell, and a *Trichoderma* cell.



More specifically, the filamentous fungal cell is an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, or *Aspergillus oryzae* cell;

5 a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium*  
10 *sporotricioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum* cell or a *Fusarium venenatum* cell (Nirenberg sp. nov; a *Humicola insolens* cell or a *Humicola lanuginosa* cell; a *Mucor miehei* cell; a *Myceliophthora thermophila* cell; a *Neurospora crassa*  
15 cell; a *Penicillium purpurogenum* cell; a *Thielavia terrestris* cell; or a *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

20 A process for preparing a filamentous fungal cell population comprising a DNA library according to the third aspect of the invention:

Placing individually different DNA sequences of interest in a  
25 filamentous fungal cell population according to step (a) of the process of the third aspect of the invention:

As state above, the specific experimental way of placing these DNA sequences within a filamentous cell may be done according to any of the many suitable techniques, such as  
30 transformation techniques. See the general fungal textbook "Fungal Genetic" (1996, ISBN 0-8247-9544-X) for a further description of such standard techniques.

A practical example may be that single stranded oligonucleotide sequences partially homologous to chromosomal  
35 DNA sequence are placed within the cell. See Calissano et al.

(Fungal genetic newsletter 43:15-16 (1995) for further description of this.

Another example may be that duplex DNA sequences comprising mismatches (e.g. comprised within a vector as shown in figure 1) are placed within the cell.

In an preferred embodiment the different DNA sequences of interest is comprised in a plasmid wherein the plasmid is characterised by that it comprises a suitable replication initiating sequence and a suitable selectable marker as described above.

Preferably the suitable replication initiating sequence is AMA1 (Gems, D., et al. (1991, Gene 98:61-67).

Growing the population of step (a) for a period of time allowing an individual DNA sequence of interest in the population to be duplicated at least once under conditions wherein the mismatch repair system gene, as describe herein has been inactivated, according to step (b) of the third aspect of the invention.

Growing of the population may be done in any of the numerous suitable known media for growing filamentous fungal cells. It is within the skilled persons general knowledge to choose such a suitable media.

As explained above an individual cell in the population must be allowed to be duplicated at least once under conditions wherein the mismatch repair system gene, as described herein has been inactivated.

Said cells may of course be allowed to be duplicated more than once under conditions wherein the mismatch repair system gene has been inactivated.

Since inactivation of the mismatch repair system normally will cause accumulation of mutations on the chromosomal DNA within the cell and thereby maybe make lethal mutations to the cell the actual preferred number of duplication cycles as described above will depend on how fast such potential lethal mutations arise.

It is within the skilled persons general knowledge to determine how many of duplication cycles it preferred.

Due to these potential lethal mutations it is preferred that the mismatch repair system under step (b) has been inactivated transitorily.

After suitable cycles of duplication according to step 5 (b) of the third aspect the transitorily inactivated mismatch repair system the mismatch repair system is then re-activated in order to avoid these lethal mutations in the filamentous fungal cell as such. The strategy for this transitorily inactivation may be any of the strategies described above.

10 Another strategy to limit introduction of mutations on the chromosome, is to transitorily stop the chromosomal replication while replicating the extra-chromosomal element under mismatch repair deficient conditions. This can be achieved by introducing mutations in elements being solely 15 necessary for the chromosomal replication.

A preferred strategy is to use a filamentous fungal cell wherein the gene involved in the mismatch repair system as described herein is permanently inactivated on the chromosome of the cell followed by inserting a plasmid into the cell which 20 comprises the gene, wherein the plasmid is characterised by that it comprises a suitable replication initiating sequence and a suitable selectable marker. See above for a further explanation of this strategy.

Preferably the suitable replication initiating sequence 25 is AMA1 (Gems, D., et al. (1991, Gene 98:61-67)).

A further embodiment relates to the process of the third aspect of the invention, wherein the mismatch repair system under step (b) is defective.

In a further embodiment the invention relates to a 30 process as described herein, wherein, under step (b) of the third aspect of the invention, there is an *in vivo* intergenic recombination of partially homologous DNA sequences of interest.

Since the overall concept of the present invention is to 35 provide a process involving inactivation of the mismatch system it is of course preferably that said partially homologous DNA sequences are capable of *in vivo* forming duplex DNA sequences comprising mismatches.

A process for production of a polypeptide of interest comprising the steps of the third aspect of the invention and wherein the DNA sequences of interest encode a polypeptide of interest, according to the fourth aspect of the invention:

Selecting from the resultant population of filamentous fungal cells of step (b) of the third aspect a desired polypeptide of interest, according to step (c) of the fourth aspect.

10       The desired polypeptide of interest may be any polypeptide comprising an desired technical feature, such as improved stability; a desired specific activity; a desired pH optimum; an improved wash performance in a detergent; etc.

15       The specific strategy for selecting this desired polypeptide of interest may be any of the numerous selecting strategies known to the skilled person, such as plate screening assays, micro-titer plate based assays, etc.

20       An embodiment of the invention relates to a process of the fourth aspect of the invention, which further comprises following steps:

25       (d) an optionally step comprising modifying the amino acid sequence of the desired selected polypeptide of interest according to a particularly further specific need;

30       (e) placing the DNA sequence encoding the polypeptide of interest of step (c) of the fourth aspect or the modified polypeptide of interest of step (d) into a filamentous fungal cell which is suitable for large scale production of the polypeptide of interest ;

      (f) cultivating the filamentous fungal cell of step (e) in a fermentor of at least 10000 m3 under conditions permitting expression of the polypeptide of interest; and

35       (g) isolating the polypeptide of interest.

      This embodiment relates to an industrial very relevant process, wherein the selected polypeptide of interest is produced in large scale.

The optionally step (d) relates to a situation wherein e.g. the desired polypeptide of interest is selected in order to e.g. identify a polypeptide with improved wash performance in a detergent according to step (c) of the third aspect of the invention. This polypeptide while having improved wash performance in a detergent may not be sufficiently stable for a commercial product. Accordingly, it may be required to make some further amino acid substitutions in this selected polypeptide, such as e.g. suitable Proline substitutions on order to obtain sufficient stability to commercialising this polypeptide.

A further embodiment relates to a process of the embodiment immediately above, wherein the filamentous fungal cell which is suitable for large-scale production of the polypeptide of interest of step (e) said embodiment is another filamentous fungal cell as compared to the filamentous fungal cell of step (a) of the third aspect of the invention.

This embodiment relates to a situation wherein the filamentous fungal cell used to select the polypeptide of interest is different from the one which is used for large scale production.

A further embodiment relates to a process as described herein, wherein the polypeptide of interest is a polypeptide derived from a filamentous fungal cell.

The term derived from a filamentous fungal cell should be understood in the sense that the information in the amino acid sequence of the polypeptide of interest is derived from a polypeptide obtained from a filamentous fungal cell.

Consequently, it may be a variant of a wild-type filamentous fungal polypeptide and/or may be a polypeptide which is a result of a recombination/shuffling of two or more different filamentous fungal polypeptides.

In an even further embodiment the invention relates to a process as described herein, wherein the polypeptide of interest is an enzyme, such as an amylase, a protease, a cellulase, a lipase, a xylanase; a phospholipase.

#### EXAMPLES:

## Materials

Chemicals used as buffers and substrates were commercial products of at least reagent grade.

5

### EXAMPLE 1:

A gel shift assay suitable for determining if a filamentous fungal cell as described herein is inactivated in the mismatch repair system:

10 The principle in this gel shift assay is that cell extracts are prepared of both (a) a filamentous fungal cell wherein the gene, as described herein, involved in the mismatch repair system is inactivated; and (b) the corresponding filamentous fungal cell wherein the gene is NOT inactivated.  
15 These extracts are then bound/mixed to oligonucleotides containing the base-pair mismatched G:T; G:A; G:G; A:C, and an extrahelical TG dinucleotide and run on a nondenaturing gel.

If the gel shift assay demonstrates that the control filamentous fungal cell wherein the gene is NOT inactivated  
20 comprises any protein(s) which binds to any of above mentioned oligonucleotides and these binding protein(s) is NOT seen in the filamentous fungal cell wherein the gene, as described herein, involved in the mismatch repair system is inactivated then it is a confirmation that the mismatch repair system in  
25 the latter is inactivated.

### Experimental procedure:

Preparation of cell extracts are performed as described in Nagata et al. (Mol. Gen Genet (1993) 237:251-260; See Materials and Methods).

30 Annealing of oligonucleotides, binding of cell extracts to duplex oligonucleotides containing mismatched, and nondenaturing polyacrylamide gelelectrophoresis are performed essentially as described (Stephenson and Karran; Selective binding to DNA base pair mismatches by proteins from human  
35 cells; J. Biol. Chem. 264:2177-21182 (1989)).

However, gelelectrophoresis is performed in TAE buffer rather than in TBE buffer. To obtain duplex oligonucleotides,

the oligonucleotide U is radiolabelled and annealed with any of the unlabelled oligonucleotides L-G.T, L-G.A, L-G.C, L-A.C, L-T.G., or L-HOM. Oligonucleotide sequences are derived from Aquilina et al. Proc. Natl. Acad. Sci. USA 91:8905-8909 (1994).

5

U: 5'-GGGAAGCTGCCAGGCCCCAGTGTTCAGCCTCCTATGCTC-3' (SEQ ID NO 3);

L-G.T: 5'-GAGCATAGGAGGCTGACATTGGGGCCTGGCAGCTTCCC-3' (SEQ ID NO 4) (resulting in a G.T mismatch);

10 L-G.A.: 5'-GAGCATAGGAGGCTGACAATGGGGCCTGGCAGCTTCCC-3' (SEQ ID NO 5) (resulting in a G.A mismatch);

L-G.G.: 5'-GAGCATAGGAGGCTGACAGTGGGGCCTGGCAGCTTCCC-3' (SEQ ID NO 6) (resulting in a G.G mismatch);

15 L-A.C.: 5'-GAGCATAGGAGGCTGACACCGGGGCCTGGCAGCTTCCC-3' (SEQ ID NO 7) (resulting in a A.C mismatch);

L-TG: 5'-GAGCATAGGAGGCTGACACTGTGGGGCCTGGCAGCTTCCC-3' (SEQ ID NO 8) (resulting in an extrahelical TG dinucleotide);

L-HOM: 5'-GAGCATAGGAGGCTGACACCGGGGCCTGGCAGCTTCCC-3' (SEQ ID NO 9) (resulting in a homoduplex).

20

In all assays, a twofold excess of unlabelled homoduplex competitor oligonucleotide is included.

#### EXAMPLE 2:

25 Cloning of a gene involved in the mismatch repair system of an *Aspergillus oryzae* cell.

The gene cloned as described in this example is shown in SEQ ID NO 1 (DNA sequence) and SEQ ID NO 2 (the translated amino acid sequence).

30 Several sequences of mismatch repair proteins from various organisms are known, only three of these have been utilized in the following: *S. cerevisiae* (M84170), *H. sapiens* (L47580) and mouse (U21011).

The numbers indicated are reference numbers from the  
35 public available GenBank database.

Based on the C-terminal homology between known mismatch repair proteins, a set of degenerate primers were designed, aiming at amplification of a partial sequence of the *Aspergillus oryzae* homolog:

5

Pr 117858 (SEQ ID NO 10): P-GGCNCARATHGGNTGYTTYGTNCC  
Pr 117859 (SEQ ID NO 11): P-GCCCANGCNARNCCRAANCC

With chromosomal DNA from *A. oryzae* strain JaL142 (WO  
10 96/29391) as template, and above primers the following 50  $\mu$ l PWO  
polymerase based PCR reaction was performed at eight MgSO<sub>4</sub>  
concentrations (0.5 mM to 4.0 mM, as recommended by the  
manufacturer; Boehringer M.). 1 mM MgSO<sub>4</sub> was found to be optimal  
and gave a discrete band of appr. 230 bp as would be expected  
15 if no introns were embodied in the sequence.

PCR-cycle profile: [96°C; 2 min - 30 cycles of (94°C; 15s -  
50°C; 15s - 72°C; 30s) - 72°C, 7 min - 4°C; hold].

20 The 230bp PCR fragment was blunt end ligated into filled  
in BamH1 site of pUC19. pUC19 was BamH1 cleaved in presence of  
calf intestine alkaline phosphatase, followed by filling in  
the sticky ends by klenow polymerase and dNTP. Three  
individual plasmids harbouring the insert were isolated from  
25 *E. coli* XL1 transformants of above ligation, and sequenced.  
Alignment of polypeptides derived by translation of the cloned  
PCR fragments, revealed a strong homolgy to known mismatch  
repair protein sequences (see Figure 2).

The underlined sequences of Figure 2 are sequences  
30 derived from the consensus PCR primers described above.

The three *Aspergillus* sequences of Figure 2 are equal to  
the sequence shown in SEQ ID NO 2 from positions 683-758,  
except from position 685 which in the final cloned sequence is  
a Thr (T) in sted of an Ile (I) as indicated above. This is due  
35 to the sequence in above mentioned consensus primers.

The alignment shown in Figure 2 clearly demonstrates that  
the cloned fragment originates from an *A. oryzae* homologue of a  
mismatch repair protein.



In order to clone the entire gene, a radiolabeled probe of the cloned fragment was generated by PCR, using 0.5 mg pUC19'msh2'-13 (see above) as template in a 100 ml reaction with Taq polymerase, 30 pmol pUC forward and reverse primers and 0.2 mM of dG-, dC-, dTTP and 0.2mM dATP + <sup>32</sup>P-dATP. The generated radiolabeled probe was liberated from pUC19 sequences by EcoR1- Hind3 digestion followed by gel purification of the resulting 293 bp fragment.

The probe was hybridized to a membrane gridded cosmid library of genomic DNA from *A. oryzae* strain A1560 (the father of JaL142) (WO96/29391). A positive clone was identified on the filter when analyzed in a phosphoimager, and the clone was identified as λ31A2.

The λ31A2 cosmid DNA was propagated and used for southern analysis, using the same radiolabeled primer as above. An approximately 9 Kb Pst fragment, split by BstX (previously found in the cloned PCR fragment) into 5.8 and 3.2 kb fragments both lightening up with the probe, was identified and cloned into Pst cut pUC19, giving a plasmid named pUC19msh2P. The insert was sequentially sequenced, starting with primers pointing out from the previously determined sequence, followed by primers based on the sequences determined in the last run:

```

130740 (SEQ ID NO 12): GCTCGAAACATCCAACATCC
25 130741 (SEQ ID NO 13): GCTGTGAATCACTTGCACC
    131928 (SEQ ID NO 14): CTTCATAAACTGCGACAAATCATGC
    131929 (SEQ ID NO 15): GGAGGAGCATCTTCGC
    131930 (SEQ ID NO 16): GGAACCTGAAGACTTTACTTCATCC
    134608 (SEQ ID NO 17): CCAGAAACTCGCTAACACC
30 134609 (SEQ ID NO 18): GTGCTTTGCGGACGC
    134610 (SEQ ID NO 19): CAGGACAGTAGGACGC
    135320 (SEQ ID NO 20): CGAGCGATGAACTCTGC
    135321 (SEQ ID NO 21): GCGTTGGTGGATTATCC
    136105 (SEQ ID NO 22): CGTTGCATCTATCATATACC
35 136106 (SEQ ID NO 23): GGTATATGATAGATGCAACGC

```

The 3825 bp sequence hereby determined (SEQ ID NO 1) was translated in the frame previously determined in the PCR

fragment. The resulting protein (SEQ ID NO 2) called Ao.MSH2 was aligned to the protein sequences of known mismatch repair proteins in Figure 3. From the alignment in shown in Figure 3 the cloned and sequenced DNA clearly encompasses the coding  
5 sequence for a homolog of yeast, man and mouse mismatch repair proteins, with one intron in the N-terminal part. The position of the intron was deduced by the standard splice rules, and constitutes the only possibility.

10 **EXAMPLE 3:**

Disruption of the gene cloned in example 1 on the chromosome of an *Aspergillus oryzae* cell:

For the disruption experiment the msh2 CDS was deleted from pUC19msh2P (see example 2) by PCR, introducing a NotI site  
15 instead. This was done by the primers:

137208 (SEQ ID NO 24): 5' P-CCGCGTCTCCAACAAGATGAATGG

137207 (SEQ ID NO 25): 5' P-CCGCTTTCTCGGGGTCATAGC

20 In a Pwo polymerase based PCR reaction with 2,5 mM MgSO<sub>4</sub> and 50 pg pUC19msh2P (conditions according to the manufacturer):

PCR-cycle profile: [96°C; 2min - 4 cycles of (94°C; 30s -  
25 52°C; 30s - 72°C; 3min) - 25 cycles of (94°C; 30s - 59°C; 30s - 72°C; 3min) - 72°C; 10min]

The resulting PCR product of appr. 8.9 Kb was isolated, ligated into pUC19, and transformed into *E. coli* XL1. From the  
30 resulting transformants pMsh2Δ was isolated, and the correctness of the new junction and its surroundings verified by sequencing [primer 138149 (SEQ ID NO 26): CCTTCCACTTTAATCCTAAGC]. (Xl1/pMsh2Δ: Lac3073).

In this construct it the *A.oryzae* pyrG (WO 96/29391) is  
35 inserted into the NotI site.

By using this construct the chromosomal gene is deleted in an *Aspergillus oryzae* cell according to standard techniques

known in the art for crossing in such a deleted gene on the chromosome by homologous recombination (Miller, B.L., et al., 1985 Mol. and Cell. Biol. 5:1714-1721).

5 **EXAMPLE 4:**

Construction of a plasmid comprising the mismatch repair gene shown in SEQ ID NO 1, the AMA1 replication initiating sequence; and the AmdS selectable marker:

The plasmid constructed as described below is highly  
10 suitable for making a filamentous fungal cell wherein the mismatch repair system may be transitorily inactivated, wherein this plasmid may be inserted into a mismatch disrupted strain of example 3 when the mismatch repair system shall be activated and deleted from the strain when the mismatch repair system  
15 shall be inactivated.

Disruption of the mismatch repair gene may cause the accumulation of new chromosomal mutations, thus such a strain might be genetically unstable. Consequently it was decided to perform the chromosomal disruption in a strain where mismatch  
20 repair gene was expressed from an extra chromosomal element readily lost when the  $\Delta$ mismatch repair phenotype was wanted.

The extra-chromosomal element was here a plasmid comprising AMA1 as replication initiating sequence and AmdS as selectable marker.

25 For this purpose the mismatch repair gene (SEQ ID NO 1) was cloned into an autonomously replicating construct harbouring one AMA1 repeat.

From pMT1505 (See example 5 below for description of pMT1505) the following fragments were isolated and ligated  
30 together:

5.16 kb NotI-[Hind3]\* + 3.515 kb [Sal]\*-BamH1 + 757bp  
BamH1- NotI

[ ]\* indicates that the site has been filled in by  
35 Klenow-polymerase and dNTP

From this ligation reaction pMT1505DHS was isolated (LaC 3212), and the mismatch repair expression cassette was introduced as a BamH1 - Mun1 fragment in the corresponding

sites in pMT1505DHS, resulting in the plasmid pAma-msh2 (LaC 3216).

*Aspergillus oryzae* JaL250 (see example 5) was transformed AmdS<sup>+</sup> with pAma-msh2, and the transformants checked for the ability to lose the amdS character when unselected (50% of the transformants), indicating the maintenance of this plasmid as extra chromosomal. (LaC3244 keep on acetamide + uridine).

#### EXAMPLE 5:

##### 10 Construction of plasmid pMT1505 used in example 4:

###### Plasmids

pMT1505: constructed as described below in Example 5

pHelp1: contains the *pyrG* gene from *A. oryzae* as a selective marker and the AMA1 sequences which enable autonomous replication in *A. nidulans* as described by Gems, D., et al. (1991. Gene 98: 61-67)

pToC68: as described in EP 0 531 372 (Novo Nordisk A/S)

###### Strains

20 JaL250: a derivative of *Aspergillus oryzae* A1560 in which the *pyrG* gene has been inactivated, as described in WO 98/01470

DH5a: an *E. coli* host cell purchased from GIBCO BRL (Life Technologies, Inc., Rockville MD)

25

pMT1466 was constructed by inserting an SphI/NarI fragment from pHelp1 into pToC68. pMT1489 was constructed by digesting pMT1466 with SphI and StuI, then religating. pMT1500 was constructed by digesting pMT1489 with AatII and NarI and ligating a linker. pMT1504 was constructed by digesting pMT1500 with NheI and religating.

pMT1505 was constructed by inserting a 2.7 kb XbaI fragment containing the *amdS* encoding gene from *A. nidulans* genomic DNA (Corrick, C.M., et al. 1987, Gene 53:63-71) into pMT1504 which had been cut with NheI.

**EXAMPLE 6:****Deletion of part of the *mshII* gene on the chromosome**

The plasmid p418MsHII (from lac3159) is cut with SalI and XhoI and treated with calf-intestinal phosphatase. In this manner part of the *mshII* gene is cut out. The large band (6400 bp) containing the vector and most of the *mshII* gene is isolated from a 1% agarose gel.

The plasmid pJal554 was constructed by ligating a SpeI/SspBI cut fragment (5330 bp) from pSO2 with a Asp718/NheI cut fragment (316 bp) from pSO2. Plasmid pJal554 is cut with SalI and a 2350 bp band which contains the *pyrG* gene is isolated on a 1% agarose gel. The 2350bp band with *pyrG* is ligated with the cut p418MsHII plasmid and transformed into *E. coli*. The right *E. coli* transformant is identified by restriction analysis and a plasmid preparation is made from this transformant.

The plasmid thus prepared is cut with EcoRI in order to linearize the plasmid before it is transformed into for example *Aspergillus oryzae* Jal250. Transformants are selected on minimal plates.

A transformant where a double crossover event has taken place is identified by making an *Aspergillus* chromosomal DNA prep followed by a PCR screen for full-length *mshII* gene using appropriate primers. A Southern blot is made using chromosomal DNA which is randomly fragmented with appropriate enzymes as well as appropriate probes for the deleted *mshII* fragment (which is not there any longer) as well as a positive control probe.

In order to determine an increased mutation frequency in the strain with an inactivated *mshII* gene, a screen for mutations in the *niaD* gene is made. This is done by growing the parent strain *Aspergillus* Jal250 and the *mshII* inactivated strain on plates.

A spore-suspension is made and aliquotes of spores are plated on chlorate-containing plate as described by Unkles et al. (S. E.Unkles, E.C.Campbell. Y.M.J.T. de Ruite Jacobs, M.

Broekhuisen, J.A. Macro, D.Carrez, R. Contreras, C.A.M.J.J. van den Hondel J.R. Kinghorn. The development of a homologous transformation system for *Aspergillus oryzae* based on nitrate assimilation pathway: A convenient and general selection system  
5 for filamentous fungal transformation. Molecular and general genetics V:218 p 99-104 (1989)).

The strain with no expression of the MshII protein will have a higher rate of niaD mutations (more chlorate resistant clones), than the control strain.

10

**EXAMPLE 7:**

Making In Vivo antisense mshII RNA to inhibit translation of mshII mRNA using the TAKA promoter to drive transcription of the Anti-sense mRNA

15 Anti-sense RNA expression is a well known way to down regulate expression of any gene in vivo (The design of Antisense RNA , Georg Sczakiel , Antisense and nucleic acid drug development V. 7 P. 439-444 (1997))

A pcr fragment is made using the oligo's:

20 000120j2 (SEQ ID NO 27): TCTGCGAATCGCTTGGATCCCGAACGCGACAACAC,  
000120j4 (SEQ ID NO 28): GAGCTCAGATCTCTTAGGTTCTGGACGAGAAGA,  
and pUC19msh2P as template. This PCR fragment contains the 5  
'end of the mshII gene including the presumed part of 5' mshII  
mRNA. Another PCR fragment is made using the oligo's:  
25 000120j3 (SEQ ID NO 29): GTTGTCGCGTTCGGGATCCAAGCGATTCGCAGAAG,  
1298-TAKA (SEQ ID NO 30): GCAAGCGCGCGCAATACATGGTGTTTTGATCAT,  
and pENI1298 as template (PCT DK99/00552).

Both PCR reactions are done using PWO polymerase according to the manufacturers manual (Boehringer-Mannheim).

30 The PCR fragments are purified using the Qiagen PCR purification kit (Qiagen). The two PCR fragments are mixed and a third PCR reaction is done with primer 1298-TAKA and 000120j4. In this manner the two PCR fragments are assembled.

The assembled PCR fragment is cut with BssHII and BglII,  
35 and purified from a 1.5 % agarose gel and ligated with pENI1298 which was cut with BssHII and Bgl II (purified from 1 % agarose

gel). The ligation mixture is transformed into *E. coli*. A DNA-prep is made of each of the resulting *E. coli* transformants. The assembled PCR fragment is sequenced to confirm that no unwanted mutations are introduced during the procedure. The correct construct contains the TAKA promoter, which drives the transcription of the *msHII* anti-sense mRNA.

The resulting plasmid is transformed into for example *Aspergillus oryzae* Jal250 along with pENI1298 as control, and transformant are selected on minimal plates. The resulting 10 transformants are isolated on minimal plates and incubated at 37°C until they sporulate.

In order to determine an increased mutation frequency in the strain, where the translation of the *msHII* mRNA is impeded due to *msHII* anti-sense RNA expression, a screen for mutations 15 in the *niaD* gene is made. A spore-suspension is prepared of the control transformants (pENI1298) and of the *msHII* Anti-sense RNA transformants, and equal amounts of spores are plated on to chlorate-containing plate as described by Unkles et al. (S. E.Unkles, E.C.Campbell. Y.M.J.T. de Ruite Jacobs, M. Broekhuizen, J.A. Macro, D.Carrez, R. Contreras, C.A.M.J.J. van den Hondel J.R. Kinghorn. The development of a homologous transformation system for *Aspergillus oryzae* based on nitrate assimilation pathway: A convenient and general selection system for filamentous fungal transformation. Molecular and general genetics V:218 p 99-104 (1989)). 25

The strain with no or low expression of the *MsHII* protein will have a higher rate of *niaD* mutations (more chlorate resistant clones), than the control strain.

## CLAIMS

1. A filamentous fungal cell, wherein a gene involved in the mismatch repair system has been inactivated and in which the gene involved in the mismatch repair system comprises:
  - 5 (a) a DNA sequence encoding the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2; or
  - (b) a DNA sequence encoding a polypeptide sequence which is at least 70% identical to the polypeptide sequence shown in positions 683-758 of SEQ ID NO 2.
- 10 2. A filamentous fungal cell, wherein a gene involved in the mismatch repair system has been inactivated and in which the gene involved in the mismatch repair system comprises:
  - 15 (a) a DNA sequence encoding the polypeptide sequence shown in positions 1-940 of SEQ ID NO 2; or
  - (b) a DNA sequence encoding a polypeptide sequence which is at least 70% identical to the polypeptide sequence shown in positions 1-940 of SEQ ID NO 2.
- 20 3. The filamentous fungal cell of claim 1 or 2, wherein the gene involved in the mismatch repair is defective.
4. The filamentous fungal cell of claim 1 or 2, wherein the gene involved in the mismatch repair has been inactivated
- 25 transitorily.
5. The filamentous fungal cell of any of the preceding claims, wherein the filamentous fungal cell is a strain of *Fusarium* or more preferably *Aspergillus*.
- 30 6. A process for preparing a filamentous fungal cell population wherein individual cells in the population comprise individually different DNA sequences of interest representing a DNA library of interest comprising following steps:
  - 35 (a) placing individually different DNA sequences of interest in a filamentous fungal cell population comprising a filamentous fungal cell of claim 1 or 2; and



(b) growing the population of (a) for a period of time allowing an individual DNA sequence of interest in the population to be duplicated at least once under conditions wherein the mismatch repair system gene of claim 1 or 2 has been inactivated.

7. The process of claim 6, wherein the mismatch repair system under step (b) is defective.

8. The process of claim 6, wherein the mismatch repair system under step (b) has been inactivated transitorily.

9. The process of any of claims 6-8, wherein the filamentous fungal cell is a strain of *Fusarium* or more preferably *Aspergillus*.

10. The process of any of claims 6-9, wherein, under step (b) of claim 6, there is an *in vivo* intergenic recombination of partially homologous DNA sequences of interest.

11. A process for production of a polypeptide of interest comprising the steps of claim 6 and wherein the DNA sequences of interest encode a polypeptide of interest and which further comprises following step:

(c) selecting from the resultant population of filamentous fungal cells of step (b) of claim 6 a desired polypeptide of interest.

12. The process of claim 11, which further comprises following steps:

(d) an optionally step comprising modifying the amino acid sequence of the desired selected polypeptide of interest according to a particularly further specific need;

(e) placing the DNA sequence encoding the polypeptide of interest of step (c) of claim 11 or the modified polypeptide of interest of step (d) into a filamentous

fungus cell which is suitable for large scale production of the polypeptide of interest;

- (f) cultivating the filamentous fungus cell of step (e) in a fermentor of at least 10000 m<sup>3</sup> under conditions permitting expression of the polypeptide of interest; and
- (g) isolating the polypeptide of interest.

13. The process of claim 12, wherein the filamentous fungus cell which is suitable for large scale production of the polypeptide of interest of step (e) of claim 12 is another filamentous fungus cell as compared to the filamentous fungus cell of step (a) of claim 6.

14. The process of any of claims 11-13, wherein the polypeptide of interest is a polypeptide derived from a filamentous fungus cell.

15. The process of any of claims 11-14, wherein the polypeptide of interest is an enzyme, such as an amylase, a protease, a cellulase, a lipase, a xylanase; a phospholipase.

1/8

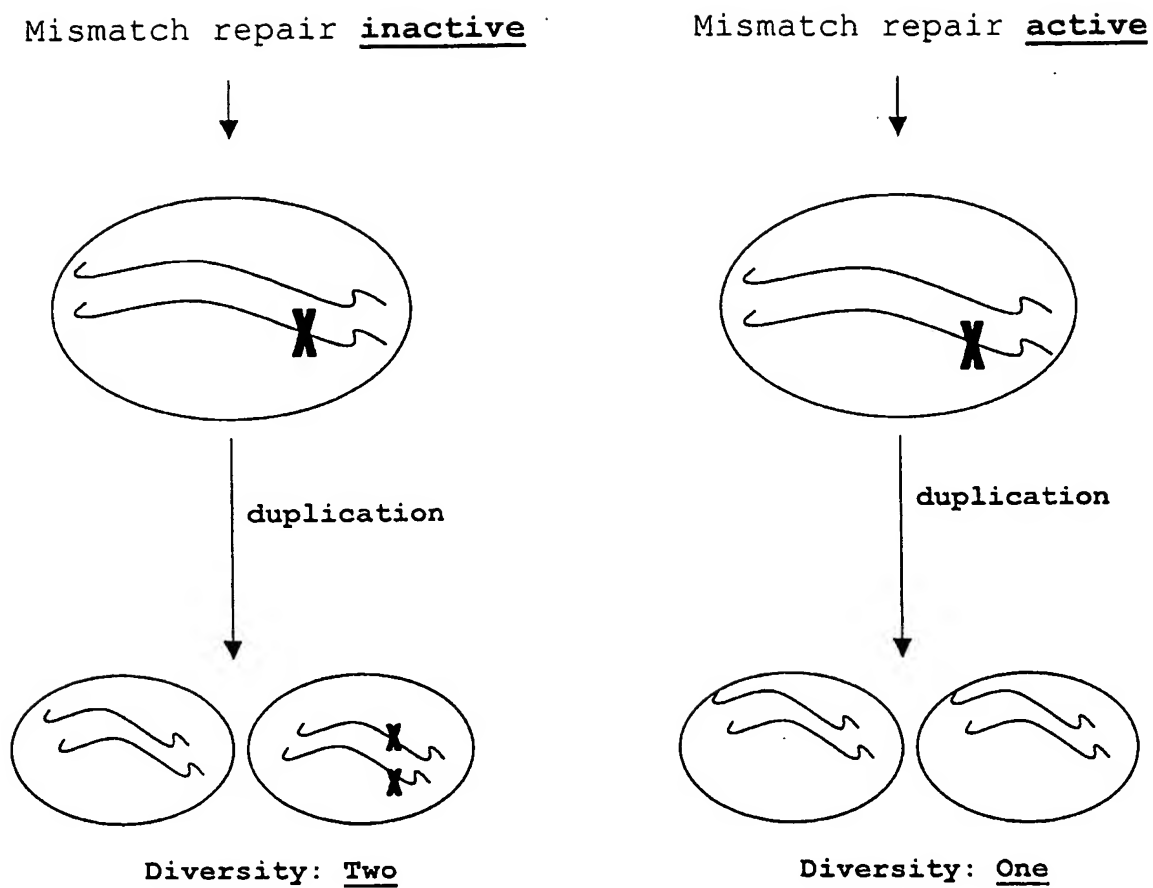


Fig. 1

70

Contig# 1	..... .   . . . . ..... . ... ..  .   .... .....
'msh2' Ao-col10	<u>AQIGCFVPCTEAELTIFDCILARVGASDSQLKGVSTFMAEMLET</u> SNILKSATSESLIIIDELGRGTSTYDGFGLAW
'msh2' Ao-col13	<u>AQIGCFVPCTEAELTIFDCILARVGASDSQLKGVSTFMAEMLET</u> SNILKSATSESLIIIDELGRGTSTYDGFGLAW
'msh2' Ao-col15	<u>AQIGCFVPCTEAELTIFDCILARVGASDSQLKGVSTFMAEMLET</u> SNILKSATSESLIIIDELGRGTSTYDGFGLAW
msh2-human.p	AQIGCFVPCESAEVSI VDCILARVGAGDSQLKGVSTFMAEMLETASILRSATKDSLIIIDELGRGTSTYDGFGLAW
S.c. msh2	AQIGCFVPCEAAEIAIVDAILCRVGAGDSQLKGVSTFMVEIILETASILKNASKNSLIIIVDELGRGTSTYDGFGLAW

2/8

Fig. 2

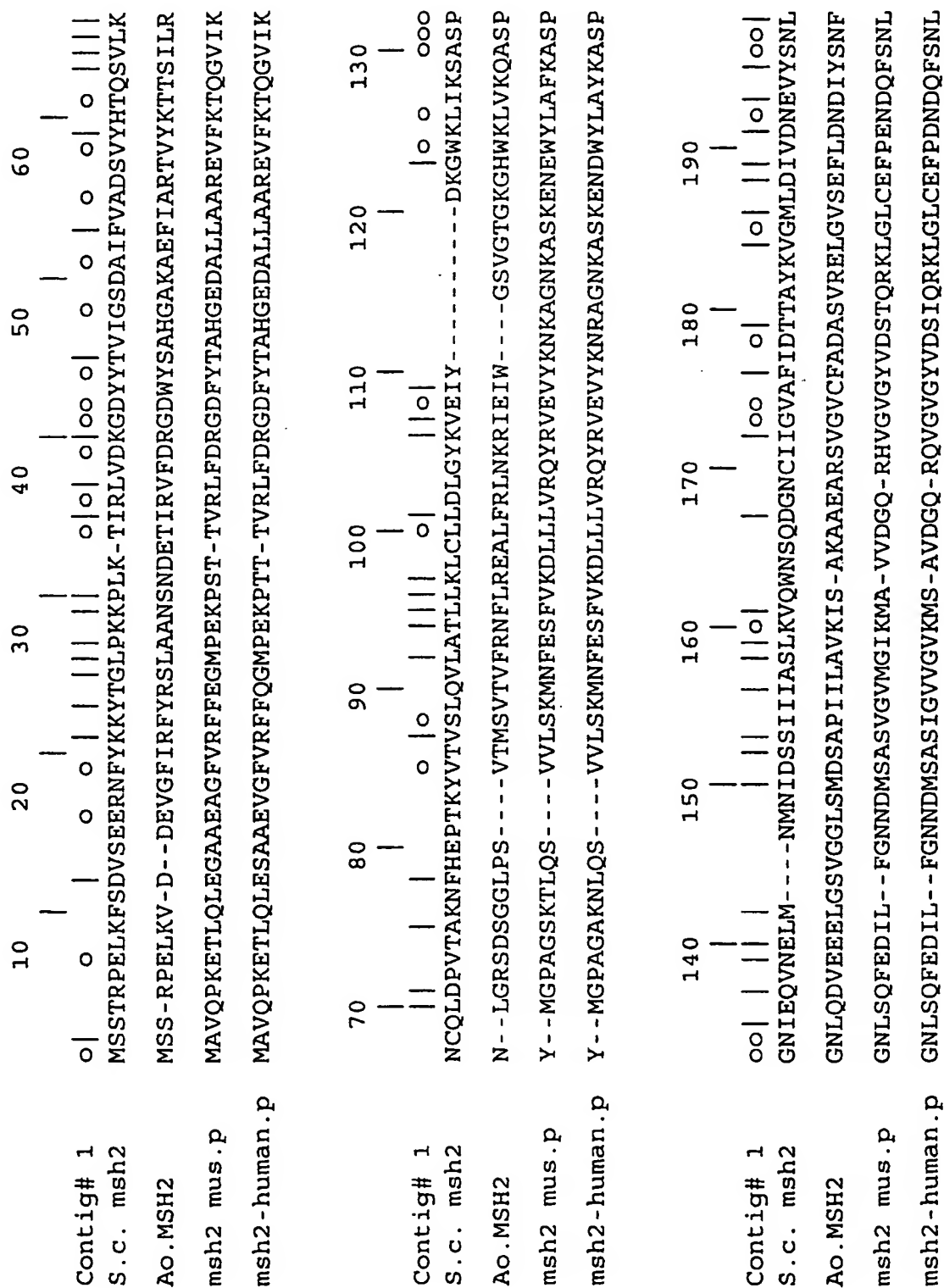


Fig. 3



	400	410	420	430	440	450	460
Contig# 1	oooo     o   oo   oo						
S.c. msh2	YLIDQIELRQMLTSEYLPMPIDIRRLTKKLNKRG-NLEDVLKIYQFSKRIPVIVQVTSFLEDDSP						
Ao.MSH2	AFVVNTELRQTMQEEHLRSIPDLYRLAKRFQRKQANLEDDVVRVYQVAIRLPGFVNLENVMDEEYQ						
msh2 mus.p	AFVEDSELRQSLQEDLLRRFPDLNRLAKKFQRQAANLQDCYRLYGINQLPSVIALEY - EGRHQ						
msh2-human.p	AFVEDAELRQTLOEDLLRRFPDLNRLAKKFQRQAANLQDCYRLYGINQLPNVIALEKH - EGKHQ						
	470	480	490	500	510	520	
Contig# 1							
S.c. msh2	TETPVELVRSVWLAPLSHHVEPLSKFEEMVETTVDLDAYEENNEFMIKVEFNELGKIRSKLDTLR						
Ao.MSH2	T-----PLETEYTSNLRSHSDSLAKLEEMVETTVDLDALE-NHEFIKPEFDESRIIRKKLDKLR						
msh2 mus.p	A-----LLLAVFTVPTLIDLRSDFSKFQEMIETTLDMDQVE-NHEFLVKPSFDPNLSELREVMGDGLE						
msh2-human.p	K-----LLLAVFTVPTLIDLRSDFSKFQEMIETTLDMDQVE-NHEFLVKPSFDPNLSELREIMNDLE						
	530	540	550	560	570	580	590
Contig# 1							
S.c. msh2	DEIHSIHLSAEDLGDFDPDKKLKENHHLHGWCMLTRNDAKELRKHKYIELSTVKAGIFFSTKQ						
Ao.MSH2	HDMGVEHRRVARLDQDIEKKLFLENHRVHGWCFLTRNESGCIRNKREYQECSTQKNGVYFTTST						
msh2 mus.p	KKMQSTLINAAARGGLDGPCKQIKLDSSAQFGYYFRVTCKEEKVLNRNNKNFSTVDIQKNGVKFTNSE						
msh2-human.p	KKMQSTLISAARDLGLDGPCKQIKLDSSAQFGYYFRVTCKEEKVLNRNNKNFSTVDIQKNGVKFTNSK						

6/8

	600	610	620	630	640	650	660
Contig# 1							
S.c. msh2	LKSIANETNILQKEYDKQQSALVREIINITLTYTPVFEKLSVLAHLDVIASFHTSSYAPIPYIR	o	o	o o o	o	o o	oooo o oo oo o
Ao.MSH2	MQTLRREHDQLSSNRYRTQTGLVNEVVNVAASCPVLERLAGVIAHLDVIVSFAHASVHAPTPYAR						
msh2 mus.p	LSSLNEEYTKNKGEYEEAQDAIVKEIVNISSGYVEPMQTLNDVLAHLDIAIVSFAHVSNAAPVPYVR						
msh2-human.p	LTSLNEEYTKNKTEYEEAQDAIVKEIVNISSGYVEPMQTLNDVLAQLDAVVVSFAHVSNGAPVPYVR						
	670	680	690	700	710	720	
Contig# 1	o		o	oo	o oo	oo ooo	ooo
S.c. msh2	PKLHPMDSEERRTHLISSRHPVLEMQDDISFISNDVTLESGKGDFLIITGPNMGGKSTYIRQVGVIS						
Ao.MSH2	PKMHPRGTGNTV-LKEARHPCMEMQDDISFITNDVALVRDESSFLIITGPNMGGKSTYIRQIGVIA						
msh2 mus.p	PVILEKGKGRII-LKASRHACVEVQDEVAFIPNDVHFEEKDKQMFHIIITGPNMGGKSTYIRQTGVIV						
msh2-human.p	PAILEKGQGRII-LKASRHACVEVQDEIAFIPNDVVFEEKDKQMFHIIITGPNMGGKSTYIRQTGVIV						
	730	740	750	760	770	780	790
Contig# 1	oooo oooooo oo  o o oo oooo oooooooooooo o ooo oo  o	oooo ooooo					
S.c. msh2	LMAQIGCFVPCCEAEIAIVDAILCRVGAGDSQLKGVSTFMVEILETASILKNASKNSLIIVDELGR						
Ao.MSH2	LMAQTGCFVPCTEAELTIFDCILARVGASDSQLKGVSTFMAEMLETSNILKSATSESLIIIDEELGR						
msh2 mus.p	LMAQIGCFVPCESAEVSIIVDCILARVGAGDSQLKGVSTFMAEMLETSSILRSATKDSLIIDEELGR						
msh2-human.p	LMAQIGCFVPCESAEVSIIVDCILARVGAGDSQLKGVSTFMAEMLETASILRSATKDSLIIDEELGR						

Fig. 3 (Continued)



7/8

	800	810	820	830	840	850
Contig# 1						
S.c. msh2	oooooooooooo	o   o   oooooooo	o     o o oo o	o  o oo o		
Ao.MSH2	GTSTYDGFGLAWAIAEHIAASKIGCFALFATHFHELTSEKLPN-VKNMHVVAHIEKNLKEQKHDD					
msh2 mus.p	GTSTYDGFGLAWAISEHIVTEIRCFGLFATHFHELTALADRYPKSVKNLHVVAFIGDGTDDDDSEDK					
msh2-human.p	GTSTYDGFGLAWAISDYIATKIGAFCMFATHFHELTALANQIP-TVNNLHVTALTTEET-----					
	860	870	880	890	900	910
Contig# 1						
S.c. msh2	o oo o o ooooooooo	oo    o oo oo	oo     o oo oo	oo     o oo oo		
Ao.MSH2	-----EDITLLYKVEPGISDQSFSGIHVAEVVQFPEKIVKMAKRKANELDDLKTNNEDLK-----KA					
msh2 mus.p	KSKRNQVTL LYRVEPGICDQSFSGIHVAELVRFPEKVVNVMARQKAELEEDFTS-SEQDQSSMAID					
msh2-human.p	-----LTMLYQVKKGVCDSQSFSGIHVAELANFPRHVIACAKQKALELEEFQNI GTSLGCDEAEPA					
	860	870	880	890	900	910
Contig# 1						
S.c. msh2	o oo o o ooooooooo	oo    o oo oo	oo     o oo oo	oo     o oo oo		
Ao.MSH2	-----EDITLLYKVEPGISDQSFSGIHVAEVVQFPEKIVKMAKRKANELDDLKTNNEDLK-----KA					
msh2 mus.p	KSKRNQVTL LYRVEPGICDQSFSGIHVAELVRFPEKVVNVMARQKAELEEDFTS-SEQDQSSMAID					
msh2-human.p	-----LTMLYQVKKGVCDSQSFSGIHVAELANFPRHVIACAKQKALELEEFQNI GTSLGCDEAEPA					

Fig. 3 (Continued)

Contig# 1	930	940	950	960	970	980	990
S.c. msh2	o o    o		o o		o		
Ao.MSH2	KLSLQEVN	EGNIRLKALLKEWIRKVEEGLHDP	SKITEEASQHKIQELLR	AIANEPEKENDNYLEI			
msh2 mus.p	KYSQEEVEEGS	-----ALLKAMLLKWKSETE	-----SSGKELTV	EEKRQIMRDLVKADEKLQANKV	FQG		
msh2-human.p	KRRCLEREQGE	-----KIILEFLSKVKQVPF	-----TAMSEESI	SAKLKQLKAEVVAKNNSFVNEI	IISR		
	KKCYLEREQGE	-----KIIQEFLSKVKQMPF	-----TEMSEENITIKLQ	LKAEVIAKNNSFVNEI	IISR		
Contig# 1	o						
S.c. msh2	YKSPCCYN						
Ao.MSH2	IK--A--L						
msh2 mus.p	IKAPA--P						
msh2-human.p	IKV-T--T						

Fig. 3 (Continued)

## SEQUENCE LISTING

&lt;110&gt; Novo Nordisk A/S

&lt;120&gt; A filamentous fungal mismatch repair defective cell

&lt;130&gt; Mismatch repair deficient fila. fungi

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 30

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 3823

&lt;212&gt; DNA

&lt;213&gt; Aspegillus oryzae

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (700)..(723)

&lt;220&gt;

&lt;221&gt; intron

&lt;222&gt; (724)..(780)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (781)..(3579)

&lt;400&gt; 1

```

gtgggtagtg gcccaaaagc tactgtggct gccagagga agccgtttcg cctagagatg 60
atcgtacaga acgttcagga ctcacgaag aatcctatct cagagaaaga ggtggaaatc 120
tgcgtcgaag tgctggctcg gcccgacatt gctggacaat gggtcgattt cgtcaccgtg 180
aatcatatca aatcggtggt tctgaaatcc tccgcggata tcaacctcaa ggatattcgg 240
gcgaagggtg gtgaactgaa gtcggcgag gacgagcctg cttcagcctc gaaccctaa 300
tcaagacttc tactgtttaa tgtgtgtttt ggatgtttgg tgttgctggg ttggataccc 360
catctgtgga gtttgatata catgacttta ttaccacct tgtgtagcat ctgttagcgt 420
tgcattatc atataccatt ttagctatta gagaatacat atcaatcatg atgaacatga 480
agtacttcag ttctatgtg taggcgtttt ttggacctat cactttgtag agttctaacg 540
ggtgctatta ttggccaatg atacttcaat atatcgcaa cgcgacaaca cgtgaccgcg 600
ttgccacgga gcgcgtctcc tggaattatc gaaatagatc gatcacggca gagtaaatgg 660
tcagtcttcc attcatcttg ttggagacta actggcaag atg tct tct cgt cca 714
                               Met Ser Ser Arg Pro
                               1           5

```

gaa ctt aag gtaagtaaac aagacaaccg gtctctcgaa cattcaataa	763
Glu Leu Lys	
caattaaccc tgttttag gtt gac gac gaa gtc ggc ttc att cgt ttt tac	813
Val Asp Asp Glu Val Gly Phe Ile Arg Phe Tyr	
10 15	
cgt tcc ctc gca gca aat agc aac gat gaa act att cgc gtt ttc gac	861
Arg Ser Leu Ala Ala Asn Ser Asn Asp Glu Thr Ile Arg Val Phe Asp	
20 25 30 35	
cgc ggt gac tgg tac tct gcc cat ggc gcc aaa gca gag ttc atc gct	909
Arg Gly Asp Trp Tyr Ser Ala His Gly Ala Lys Ala Glu Phe Ile Ala	
40 45 50	
cgc act gtg tac aag acc acc tct ata ctc cgc aat cta ggt cgc agc	957
Arg Thr Val Tyr Lys Thr Thr Ser Ile Leu Arg Asn Leu Gly Arg Ser	
55 60 65	
gac tca gga ggc ctt ccc tcc gtc acc atg agt gtc acc gtc ttc cgt	1005
Asp Ser Gly Gly Leu Pro Ser Val Thr Met Ser Val Thr Val Phe Arg	
70 75 80	
aac ttt ctc cgc gaa gct ctc ttc cga ctc aac aag cgc att gaa atc	1053
Asn Phe Leu Arg Glu Ala Leu Phe Arg Leu Asn Lys Arg Ile Glu Ile	
85 90 95	
tgg ggc tca gtc gga acg ggc aag ggt cat tgg aag ctg gta aag caa	1101
Trp Gly Ser Val Gly Thr Gly Lys Gly His Trp Lys Leu Val Lys Gln	
100 105 110 115	
gct agc ccg gga aac ctc caa gat gtg gaa gaa gag ttg ggc agc gtt	1149
Ala Ser Pro Gly Asn Leu Gln Asp Val Glu Glu Glu Leu Gly Ser Val	
120 125 130	
ggt gga tta tcc atg gac tcg gct cca att atc cta gca gtg aag atc	1197
Gly Gly Leu Ser Met Asp Ser Ala Pro Ile Ile Leu Ala Val Lys Ile	
135 140 145	
tcg gcc aag gcc gca gag gct agg agt gtg gga gtg tgc ttt gcg gac	1245
Ser Ala Lys Ala Ala Glu Ala Arg Ser Val Gly Val Cys Phe Ala Asp	
150 155 160	
gca agt gta cgg gaa ctc ggt gtt agc gag ttt ctg gat aac gat atc	1293
Ala Ser Val Arg Glu Leu Gly Val Ser Glu Phe Leu Asp Asn Asp Ile	
165 170 175	
tat tcc aac ttt gag tcg ctt att atc caa ctc ggg gtg aag gag tgt	1341
Tyr Ser Asn Phe Glu Ser Leu Ile Ile Gln Leu Gly Val Lys Glu Cys	
180 185 190 195	
ttg gtg cag atg gat gct aat aag aag gat gtt gag ctg gga aag att	1389
Leu Val Gln Met Asp Ala Asn Lys Lys Asp Val Glu Leu Gly Lys Ile	
200 205 210	
cgg gct att gcg gat agt tgt ggg atc gct atc tcc gag agg ccg gtg	1437
Arg Ala Ile Ala Asp Ser Cys Gly Ile Ala Ile Ser Glu Arg Pro Val	
215 220 225	

gct gat tat ggt gtc aag gat att gag cag gat ctg acg agg ttg ttg	1485
Ala Asp Tyr Gly Val Lys Asp Ile Glu Gln Asp Leu Thr Arg Leu Leu	
230 235 240	
agg gat gaa cgg tcg gct ggt acg ctg ccg cag acg gag cta aag ctt	1533
Arg Asp Glu Arg Ser Ala Gly Thr Leu Pro Gln Thr Glu Leu Lys Leu	
245 250 255	
gcg atg gcc tcg ccg tct gcg ttg atc aag tac ctt ggg gtt atg acg	1581
Ala Met Gly Ser Ala Ser Ala Leu Ile Lys Tyr Leu Gly Val Met Thr	
260 265 270 275	
gat cct aca aac ttc ggc cag tac cag ctg tat cag cat gat ttg tcg	1629
Asp Pro Thr Asn Phe Gly Gln Tyr Gln Leu Tyr Gln His Asp Leu Ser	
280 285 290	
cag ttt atg aag ttg gat tcg tcg gcg ctg cgt gct ctt aac ctt atg	1677
Gln Phe Met Lys Leu Asp Ser Ser Ala Leu Arg Ala Leu Asn Leu Met	
295 300 305	
cct ggt ccg cgg gac gga tcg aag tct atg agt ttg ttt ggt ttg ttg	1725
Pro Gly Pro Arg Asp Gly Ser Lys Ser Met Ser Leu Phe Gly Leu Leu	
310 315 320	
aat cac tgc aag acc cct gtt ggt agc cgg ttg ctt gcg cag tgg ctg	1773
Asn His Cys Lys Thr Pro Val Gly Ser Arg Leu Leu Ala Gln Trp Leu	
325 330 335	
aaa cag ccg ttg atg gat ctg gcg gag atc gag aag aga cag cag ctt	1821
Lys Gln Pro Leu Met Asp Leu Ala Glu Ile Glu Lys Arg Gln Gln Leu	
340 345 350 355	
gtt gag gcg ttt gtt gtt aac acg gag ctg aga cag act atg cag gag	1869
Val Glu Ala Phe Val Val Asn Thr Glu Leu Arg Gln Thr Met Gln Glu	
360 365 370	
gag cat ctt cgc tcc ata ccg gat ctg tat aga cta gcg aag cgg ttc	1917
Glu His Leu Arg Ser Ile Pro Asp Leu Tyr Arg Leu Ala Lys Arg Phe	
375 380 385	
cag cgc aaa cag gca aac ttg gaa gac gtt gtg cgg gtg tac cag gtt	1965
Gln Arg Lys Gln Ala Asn Leu Glu Asp Val Val Arg Val Tyr Gln Val	
390 395 400	
gct att cgt ttg cct ggt ttt gtc aac tct ctg gag aat gtt atg gat	2013
Ala Ile Arg Leu Pro Gly Phe Val Asn Ser Leu Glu Asn Val Met Asp	
405 410 415	
gaa gag tat cag acg ccc ctg gag acg gag tat act tcc aac ctg cgg	2061
Glu Glu Tyr Gln Thr Pro Leu Glu Thr Glu Tyr Thr Ser Asn Leu Arg	
420 425 430 435	
agt cac tct gat agc tta gcg aaa ctg gag gag atg gtt gag act acg	2109
Ser His Ser Asp Ser Leu Ala Lys Leu Glu Glu Met Val Glu Thr Thr	
440 445 450	
gtt gac ctt gat gcg ctg gag aac cac gag ttc atc atc aag cct gag	2157
Val Asp Leu Asp Ala Leu Glu Asn His Glu Phe Ile Ile Lys Pro Glu	
455 460 465	

ttt gac gag agt ctg cgg atc atc agg aag aag ctg gac aag ctc cgt Phe Asp Glu Ser Leu Arg Ile Ile Arg Lys Lys Leu Asp Lys Leu Arg 470 475 480	2205
cat gat atg ggc gtt gag cac cgc agg gta gct cgg gac ctt gac caa His Asp Met Gly Val Glu His Arg Arg Val Ala Arg Asp Leu Asp Gln 485 490 495	2253
gat att gag aag aag ttg ttc ctg gag aac cac agg gtg cac gga tgg Asp Ile Glu Lys Lys Leu Phe Leu Glu Asn His Arg Val His Gly Trp 500 505 510 515	2301
tgc ttc cga ctt act cgc aac gag tcg gga tgc atc cgc aat aag aga Cys Phe Arg Leu Thr Arg Asn Glu Ser Gly Cys Ile Arg Asn Lys Arg 520 525 530	2349
gag tac cag gaa tgt tct aca cag aag aac ggt gtc tac ttc act acg Glu Tyr Gln Glu Cys Ser Thr Gln Lys Asn Gly Val Tyr Phe Thr Thr 535 540 545	2397
tcg act atg caa acc ttg cgc cgg gag cat gat caa ctg tcc tcg aac Ser Thr Met Gln Thr Leu Arg Arg Glu His Asp Gln Leu Ser Ser Asn 550 555 560	2445
tac aat aga act cag acc ggc ctg gtg aat gag gtc gtt aac gtt gcc Tyr Asn Arg Thr Gln Thr Gly Leu Val Asn Glu Val Val Asn Val Ala 565 570 575	2493
gcg tcc tac tgt cct gtt ttg gaa cga ctt gcc ggt gtc ata gca cac Ala Ser Tyr Cys Pro Val Leu Glu Arg Leu Ala Gly Val Ile Ala His 580 585 590 595	2541
ctc gat gtc att gta agc ttc gct cat gct tct gtt cat gcg ccg acc Leu Asp Val Ile Val Ser Phe Ala His Ala Ser Val His Ala Pro Thr 600 605 610	2589
ccc tat gct cgg ccc aag atg cac cca cga ggc acc gga aac aca gtt Pro Tyr Ala Arg Pro Lys Met His Pro Arg Gly Thr Gly Asn Thr Val 615 620 625	2637
ctc aag gaa gcg cgc cac ccc tgt atg gaa atg cag gat gat att tca Leu Lys Glu Ala Arg His Pro Cys Met Glu Met Gln Asp Asp Ile Ser 630 635 640	2685
ttc att act aat gat gtg gct ttg gtc cga gac gag tcc tcc ttc ctc Phe Ile Thr Asn Asp Val Ala Leu Val Arg Asp Glu Ser Ser Phe Leu 645 650 655	2733
atc att act ggt cct aac atg gga ggt aaa tcg act tat att cgc caa Ile Ile Thr Gly Pro Asn Met Gly Gly Lys Ser Thr Tyr Ile Arg Gln 660 665 670 675	2781
att ggt gtt atc gct ctc atg gct cag acg ggc tgc ttt gtg cct tgt Ile Gly Val Ile Ala Leu Met Ala Gln Thr Gly Cys Phe Val Pro Cys 680 685 690	2829
aca gaa gca gaa ttg acc atc ttt gac tgt atc ctt gca cgt gtt ggt Thr Glu Ala Glu Leu Thr Ile Phe Asp Cys Ile Leu Ala Arg Val Gly 700 705 710 715	2877

695			700			705			
gca agt gat tca cag ctc aag gga gtt tcc act ttc atg gct gag atg	2925								
Ala Ser Asp Ser Gln Leu Lys Gly Val Ser Thr Phe Met Ala Glu Met									
	715							720	
ctc gaa aca tcc aac atc ctc aag tcg gca acg tcc gag tct ctt atc	2973								
Leu Glu Thr Ser Asn Ile Leu Lys Ser Ala Thr Ser Glu Ser Leu Ile									
	725							735	
atc atc gac gag ctt ggg cgc ggt aca agc acg tat gac gga ttc ggc	3021								
Ile Ile Asp Glu Leu Gly Arg Gly Thr Ser Thr Tyr Asp Gly Phe Gly									
	740							750	755
cta gca tgg gcc atc tct gaa cac atc gtc aca gag att cgt tgc ttc	3069								
Leu Ala Trp Ala Ile Ser Glu His Ile Val Thr Glu Ile Arg Cys Phe									
	760							765	770
ggc ctt ttc gcc act cac ttc cat gaa ttg aca gct ctc gcc gat cga	3117								
Gly Leu Phe Ala Thr His Phe His Glu Leu Thr Ala Leu Ala Asp Arg									
	775							780	785
tac ccc aag tct gtc aag aac ctg cac gta gtc gcc ttc atc ggc gat	3165								
Tyr Pro Lys Ser Val Lys Asn Leu His Val Val Ala Phe Ile Gly Asp									
	790							800	
ggt act gat gat gac agt gaa gat aag aag tcc aag cgg aac cag gtc	3213								
Gly Thr Asp Asp Asp Ser Glu Asp Lys Lys Ser Lys Arg Asn Gln Val									
	805							810	815
act ctt ctg tac cgg gtc gaa cct ggc att tgt gac cag tca ttc ggt	3261								
Thr Leu Leu Tyr Arg Val Glu Pro Gly Ile Cys Asp Gln Ser Phe Gly									
	820							825	830
atc cac gtt gcc gaa ttg gtc cgc ttc ccg gag aag gtg gtc aac atg	3309								
Ile His Val Ala Glu Leu Val Arg Phe Pro Glu Lys Val Val Asn Met									
	840							845	850
gcc cgc cag aag gca gag gaa ctt gaa gac ttt act tca tcc gaa cag	3357								
Ala Arg Gln Lys Ala Glu Glu Leu Glu Asp Phe Thr Ser Ser Glu Gln									
	855							860	865
caa gac cag cag tca tcc atg gcg atc gat aaa tac tcg cag gaa gaa	3405								
Gln Asp Gln Gln Ser Ser Met Ala Ile Asp Lys Tyr Ser Gln Glu Glu									
	870							875	880
gtt gag gag ggc agt gcc ctt ctc aaa gcg atg ctg ctg aaa tgg aag	3453								
Val Glu Glu Gly Ser Ala Leu Leu Lys Ala Met Leu Leu Lys Trp Lys									
	885							890	895
tcg gag acc gag tcc tct ggt aag gag ttg aca gtg gaa gag aag cga	3501								
Ser Glu Thr Glu Ser Ser Gly Lys Glu Leu Thr Val Glu Glu Lys Arg									
	900							905	910
cag atc atg cgt gat ctc gtc aaa gca gat gag aag ctg caa gca aac	3549								
Gln Ile Met Arg Asp Leu Val Lys Ala Asp Glu Lys Leu Gln Ala Asn									
	920							925	930
aag gtc ttc cag ggt atc aag gct tta tag attagtagttt gcgtcttttt	3599								

Lys Val Phe Gln Gly Ile Lys Ala Leu  
935 940

tctttctcgg ggtcatagcg gttcggcggtt tggaaaggtgt caatctgtgt atgtgtgac 3659  
tacggacatg aggataaaat gtgtagggaa taatattatc caaaaatttt cgagtgtattg 3719  
cttcttttggga catatcggtt aggattaaag tggaaaggga gaaatcccat tcaactatat 3779  
cgacataagt cacgttgaga tcgcgagtcct agacgctcac cggg 3823

<210> 2

<211> 940

<212> PRT

<213> *Aspegillus cryzae*

<400> 2

Met	Ser	Ser	Arg	Pro	Glu	Leu	Lys	Val	Asp	Asp	Glu	Val	Gly	Phe	Ile
1				5					10					15	
Arg	Phe	Tyr	Arg	Ser	Leu	Ala	Ala	Asn	Ser	Asn	Asp	Glu	Thr	Ile	Arg
			20					25					30		
Val	Phe	Asp	Arg	Gly	Asp	Trp	Tyr	Ser	Ala	His	Gly	Ala	Lys	Ala	Glu
		35					40					45			
Phe	Ile	Ala	Arg	Thr	Val	Tyr	Lys	Thr	Thr	Ser	Ile	Leu	Arg	Asn	Leu
	50					55					60				
Gly	Arg	Ser	Asp	Ser	Gly	Gly	Leu	Pro	Ser	Val	Thr	Met	Ser	Val	Thr
	65				70					75					80
Val	Phe	Arg	Asn	Phe	Leu	Arg	Glu	Ala	Leu	Phe	Arg	Leu	Asn	Lys	Arg
			85						90					95	
Ile	Glu	Ile	Trp	Gly	Ser	Val	Gly	Thr	Gly	Lys	Gly	His	Trp	Lys	Leu
		100						105					110		
Val	Lys	Gln	Ala	Ser	Pro	Gly	Asn	Leu	Gln	Asp	Val	Glu	Glu	Glu	Leu
		115					120					125			
Gly	Ser	Val	Gly	Gly	Leu	Ser	Met	Asp	Ser	Ala	Pro	Ile	Ile	Leu	Ala
	130					135					140				
Val	Lys	Ile	Ser	Ala	Lys	Ala	Ala	Glu	Ala	Arg	Ser	Val	Gly	Val	Cys
	145				150					155					160
Phe	Ala	Asp	Ala	Ser	Val	Arg	Glu	Leu	Gly	Val	Ser	Glu	Phe	Leu	Asp
			165						170					175	
Asn	Asp	Ile	Tyr	Ser	Asn	Phe	Glu	Ser	Leu	Ile	Ile	Gln	Leu	Gly	Val
		180						185					190		
Lys	Glu	Cys	Leu	Val	Gln	Met	Asp	Ala	Asn	Lys	Lys	Asp	Val	Glu	Leu
		195					200					205			
Gly	Lys	Ile	Arg	Ala	Ile	Ala	Asp	Ser	Cys	Gly	Ile	Ala	Ile	Ser	Glu
	210					215					220				
Arg	Pro	Val	Ala	Asp	Tyr	Gly	Val	Lys	Asp	Ile	Glu	Gln	Asp	Leu	Thr
	225				230					235					240
Arg	Leu	Leu	Arg	Asp	Glu	Arg	Ser	Ala	Gly	Thr	Leu	Pro	Gln	Thr	Glu
			245						250					255	
Leu	Lys	Leu	Ala	Met	Gly	Ser	Ala	Ser	Ala	Leu	Ile	Lys	Tyr	Leu	Gly
			260					265					270		
Val	Met	Thr	Asp	Pro	Thr	Asn	Phe	Gly	Gln	Tyr	Gln	Leu	Tyr	Gln	His
		275					280					285			
Asp	Leu	Ser	Gln	Phe	Met	Lys	Leu	Asp	Ser	Ser	Ala	Leu	Arg	Ala	Leu
	290					295					300				
Asn	Leu	Met	Pro	Gly	Pro	Arg	Asp	Gly	Ser	Lys	Ser	Met	Ser	Leu	Phe
	305				310					315					320
Gly	Leu	Leu	Asn	His	Cys	Lys	Thr	Pro	Val	Gly	Ser	Arg	Leu	Leu	Ala



[illegible]

Ile	Gly	Asp	Gly	Thr	Asp	Asp	Asp	Ser	Glu	Asp	Lys	Lys	Ser	Lys	Arg
				805					810					815	
Asn	Gln	Val	Thr	Leu	Leu	Tyr	Arg	Val	Glu	Pro	Gly	Ile	Cys	Asp	Gln
			820					825					830		
Ser	Phe	Gly	Ile	His	Val	Ala	Glu	Leu	Val	Arg	Phe	Pro	Glu	Lys	Val
		835					840					845			
Val	Asn	Met	Ala	Arg	Gln	Lys	Ala	Glu	Glu	Leu	Glu	Asp	Phe	Thr	Ser
		850				855					860				
Ser	Glu	Gln	Gln	Asp	Gln	Gln	Ser	Ser	Met	Ala	Ile	Asp	Lys	Tyr	Ser
865					870					875					880
Gln	Glu	Glu	Val	Gln	Glu	Gly	Ser	Ala	Leu	Leu	Lys	Ala	Met	Leu	Leu
			885						890					895	
Lys	Trp	Lys	Ser	Gln	Thr	Glu	Ser	Ser	Gly	Lys	Glu	Leu	Thr	Val	Glu
			900					905					910		
Glu	Lys	Arg	Gln	Ile	Met	Arg	Asp	Leu	Val	Lys	Ala	Asp	Glu	Lys	Leu
		915					920					925			
Gln	Ala	Asn	Lys	Val	Phe	Gln	Gly	Ile	Lys	Ala	Leu				
		930				935					940				

```
<210> 3
<211> 38
<212> DNA
<213> Artificial Sequence
```

<220>  
<223> Description of Artificial Sequence: oligo # U

<400> 3  
gggaagctgc caggccccag tgtcagcttc ctatqctc 38

```
<210> 4
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: oligo # L-G.T
```

<400> 4  
gagcatagga ggctgacatt ggggcctggc agcttccc 38

```
<210> 5
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: oligo #
      L-G.A.
```

```
<400> 5
gagcatagga ggctgacaat ggggectggc agcttccc 38
```

<210> 6

<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo #  
L-G.G.

<400> 6  
gagcatagga ggctgacagt ggggcctggc agcttccc 38

<210> 7  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo #  
L-A.C.

<400> 7  
gagcatagga ggctgacacc ggggcctggc agcttccc 38

<210> 8  
<211> 40  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # L-TG

<400> 8  
gagcatagga ggctgacact gtggggcctg gcagcttccc 40

<210> 9  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # L-HOM

<400> 9  
gagcatagga ggctgacacc ggggcctggc agcttccc 38

<210> 10  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # PR  
117858

<400> 10

ggcncarath ggntgyttyg tncc

24

<210> 11  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo # Pr  
117859

<400> 11  
gcccangcna rnccraancc

20

<210> 12  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo # 130740

<400> 12  
gctcgaaaca tccaacatcc

20

<210> 13  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo # 130741

<400> 13  
gctgtgaatc acttgcacc

19

<210> 14  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo # 131928

<400> 14  
cttcataaac tgcgacaaat catgc

25

<210> 15  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo # 131929

<400> 15  
ggaggagcat ctgcgc

16

<210> 16  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # 131930

<400> 16  
ggaacttgaa gactttactt catcc

25

<210> 17  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # 134608

<400> 17  
ccagaaactc gctaacacc

19

<210> 18  
<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo #  
134609

<400> 18  
gtgctttgcg gacgc

15

<210> 19  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # 134610

<400> 19  
caggacagta ggacgc

16

<210> 20  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: oligo # 135320  
  
<400> 20  
cgagcgatga acttggc 17  
  
<210> 21  
<211> 17  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: oligo # 135321  
  
<400> 21  
gcgttggtgg attatcc 17  
  
<210> 22  
<211> 20  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: oligo # 136105  
  
<400> 22  
cgttgcatct atcatatacc 20  
  
<210> 23  
<211> 21  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: oligo #  
136106  
  
<400> 23  
ggtatatgat agatgcaacg c 21  
  
<210> 24  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: oligo #  
137208  
  
<400> 24  
ccgcgtctcc aacaagatga atgg 24  
  
<210> 25  
<211> 21

<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo # 137207

<400> 25

ccgctttctc ggggcacatag c

21

<210> 26

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer 138149

<400> 26

cctttccact ttaatcctaa gc

22

<210> 27

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo #  
000120j2

<400> 27

tctgcgaatc gcttgatcc cgaacgcgac aacac

35

<210> 28

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo #  
000120j4

<400> 28

gagctcagat ctcttaggtt ctggacgaga aga

33

<210> 29

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo #  
000120j3

<400> 29

gttgatcgct tcgggatcca agcgattcgc agaag

35

<210> 30

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: oligo #  
1298-TANA

<400> 30

gcaagcgcgc gcaatacatg gtgttttgat cat

33



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/DK 00/00063

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C12N 1/15, C12N 15/10, C12P 9/00 // (C12N 1/15, C12R 1:66), (C12N 1/15, C12R 1:77)

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 9831837 A1 (MAXYGEN INC.), 23 July 1998 (23.07.98), see page 21, line 31 - page 23, line 20 and page 45, line 12 - page 56, line 21, esp. page 56, line 8 - line 12 --	1-15
A	Database SWISS-PROT, accession no. 013396, Huber D.H.: "DNA Mismatch Repair Protein MSH2. Neurospora crassa." 15 DEC 1998 --	1-15
A	WO 9705268 A1 (SETRATECH), 13 February 1997 (13.02.97), cited in the application --	1-15

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 June 2000

Date of mailing of the international search report

16-06-2000

Name and mailing address of the ISA/

Swedish Patent Office

Box 5055, S-102 42 STOCKHOLM

Facsimile No. +46 8 666 02 86

Authorized officer

Hampus Rystedt/E1s

Telephone No. +46 8 782 25 00

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 00/00063

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 9737011 A1 (SETRATECH S.A.R.L.), 9 October 1997 (09.10.97), cited in the application  --	1-15
A	WO 9007576 A1 (SETRATECH), 12 July 1990 (12.07.90), cited in the application  -- -----	1-15

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

PCT/DK 00/00063

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9831837 A1	23/07/98	AU 5920998 A	07/08/98
WO 9705268 A1	13/02/97	AU 4784996 A	26/02/97
		EP 0842289 A	20/05/98
WO 9737011 A1	09/10/97	NONE	
WO 9007576 A1	12/07/90	AT 127519 T	15/09/95
		AU 4834390 A	01/08/90
		CA 2006549 A	26/06/90
		DE 68924174 D,T	14/03/96
		EP 0449923 A,B	09/10/91
		SE 0449923 T3	
		ES 2077058 T	16/11/95
		FR 2641793 A,B	20/07/90
		IE 72469 B	09/04/97
		IL 92856 D	00/00/00
		JP 4503601 T	02/07/92
		US 5912119 A	15/06/99
		US 5965415 A	12/10/99